



Multifidelity surrogate modeling based on Radial Basis Functions

Cédric Durantin, Justin Rouxel, Jean-Antoine Desideri, Alain Glière

► To cite this version:

Cédric Durantin, Justin Rouxel, Jean-Antoine Desideri, Alain Glière. Multifidelity surrogate modeling based on Radial Basis Functions. Structural and Multidisciplinary Optimization, 2017, 56 (5), pp.1061-1075. 10.1007/s00158-017-1703-7 . hal-01660796

HAL Id: hal-01660796

<https://inria.hal.science/hal-01660796>

Submitted on 11 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Multifidelity surrogate modeling based on Radial Basis Functions

Cédric Durantin · Justin Rouxel · Jean-Antoine Désidéri · Alain Glière

Received: date / Accepted: date

Abstract Multiple models of a physical phenomenon are sometimes available with different levels of approximation. The high fidelity model is more computationally demanding than the coarse approximation. In this context, including information from the lower fidelity model to build a surrogate model is desirable. Here, the study focuses on the design of a miniaturized photoacoustic gas sensor which involves two numerical models. First, a multifidelity metamodeling method based on Radial Basis Function, the co-RBF, is proposed. This surrogate model is compared with the classical co-kriging method on two analytical benchmarks and on the photoacoustic gas sensor. Then an extension to the multifidelity framework of an already existing RBF-based optimization algorithm is applied to optimize the sensor efficiency. The co-RBF method brings promising results on a problem in larger dimension and can be considered as an alternative to co-kriging for multifidelity metamodeling.

Keywords Metamodel · RBF · Multifidelity codes · Co-RBF · Optimization

1 Introduction

In the micro-electronics industry, as well as in other fields, products are nowadays usually designed using computer models. Physical experimentation with prototypes are too costly compared to computer simula-

tions for the development of new concepts. However, numerical models are often built with complex mathematical codes involving multiple disciplines that are very expensive to evaluate. For example a single Finite Element analysis of a gas sensor can require a computational time of several hours.

A solution that has attracted intensive attention during the last decades is to replace the physics based computer simulation by a metamodel, which emulates the statistical input-output relationship. The advantage of using a metamodel is the reduction of the computational cost necessary to approximate the output of the numerical model. The methodology implies that a sampling plan is evaluated first, in order to train the metamodel. Then, the response surface is updated using an adaptive sampling algorithm, in order to reduce the prediction uncertainty at new points. Surrogate models are useful for sensitivity analysis or optimization purpose because they limit the cost of function calls compared to a direct evaluation of the simulation. The reader is referred to Wang and Shan (2007) for a non-exhaustive review of different metamodeling methods.

Kriging-based surrogate models (Kriging 1951; Santner et al 2003) are mainstream in the metamodeling community because of their accuracy, the availability of the prediction variance and their use in sequential design and optimization (Schonlau and Welch 1996). Sometimes multiple models of a physical phenomenon are available with different levels of approximation, the high fidelity model being more expensive in terms of computational time than the coarse approximation. Thus, Kriging models have recently been improved to take advantage of the use of both the coarse and the precise versions of the computer simulation, resulting in an enhanced prediction accuracy at a reduced computational cost (Kennedy and O'Hagan 2000). Then, several

Cédric Durantin · Justin Rouxel · Alain Glière
University Grenoble Alpes, CEA, LETI, MINATEC Campus,
38054 Grenoble, France
E-mail: cedric.durantin@free.fr

Jean Antoine Désidéri
Acumes Team, INRIA Sophia Antipolis Méditerranée, 06902,
Sophia Antipolis, France

studies dealing with co-kriging (multifidelity version of Kriging) and its use in optimization have been exposed (Forrester et al 2007; Dong et al 2015; Le Gratiet and Cannamela 2015).

Another possible choice in this context is the Radial Basis Functions (RBF) based interpolation method (Powell 1987; Dyn et al 1986). Among metamodeling approaches, it is one of the most effective multidimensional approximation methods (Jin et al 2001), as the dimension of the input space does not alter its performance (Powell 2001). It is also suitable for optimization since an adaptive sampling strategy has been developed by Gutmann (2001) to sequentially enrich the training database of the RBF metamodel, in order to find an optimal design of the model simulated.

The present work focuses on the optimization of a photoacoustic gas sensor using two different simulation models based on more or less sophisticated physical approaches. A possible way to solve this problem involves the use of RBF multifidelity metamodeling as an alternative to co-kriging. The use of RBF in multifidelity framework has already been proposed (Sun et al 2010), but the methodology described here is different. The optimization problem of the photoacoustic gas sensor is explained in section 2, where the reason for the choice of RBF metamodeling instead of kriging is also discussed. The RBF metamodel is described in Section 3. The new co-RBF methodology (multifidelity version of RBF metamodeling) is detailed in Section 4, followed by a proposition of an extension of the Gutmann optimization algorithm to the multifidelity framework in Section 5. A comparison between co-kriging and co-RBF metamodels is presented in Section 6, where two different analytical test problems are solved. The efficiency of the new multifidelity surrogate model is then confirmed on the design optimization of a photoacoustic gas sensor in Section 7.

2 Miniaturized photoacoustic gas sensor test problem

The design of a miniaturized photoacoustic gas sensor is the principal motivation of the methodology proposed. The physical behaviour of this component and the optimization problem are explained in the following section.

2.1 Theory and models

Photoacoustic (PA) spectroscopy is employed to detect gas traces with a high sensitivity, sometimes below the part per billion by volume (ppbv) level. The principle of PA spectroscopy relies on the excitation of a molecule

of interest by a light source emitting at the wavelength of an absorption line of the molecule. The light source, usually a laser in the mid-infrared range, is modulated at the acoustic frequency of a resonant cell, containing the gas mixture. During the molecules collisional relaxation, the kinetic energy exchange with the surrounding gas creates local temperature modulation, and thus acoustic waves in the chamber (Miklós et al 2001). The cell used in this work is of the differential Helmholtz resonator (DHR) type (Zéninari et al 1999). It is composed of two chambers linked by two capillaries (Figure 1). The gas excitation is ensured by illuminating one chamber with a laser source. At the Helmholtz resonance of the cavity, acoustic signals in the two chambers are in phase opposition. The signals provided by two microphones measuring the pressure into each chamber are subtracted to provide the PA signal. As it is inversely proportional to the volume of the resonant cell (Miklós et al 2001), an effort for reducing the size of PA cells has been initiated during the last decade (Holthoff et al 2010; Bauer et al 2014; Rouxel et al 2016).

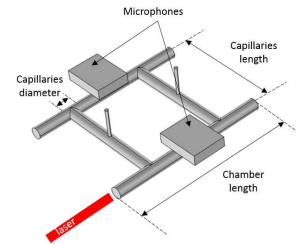


Fig. 1 Shape of the cavity that contain the gas. The cell is composed of two chambers linked by two capillaries, and two microphones cavities. The laser is illuminating one of the chamber.

Assuming no viscous and thermal losses and a harmonic heat source, the non homogeneous Helmholtz equation (1) can be used to compute the pressure field in the cell, and thus the differential PA signal:

$$\nabla^2 p + k^2 p = i\omega \frac{\gamma - 1}{c^2} \mathcal{H} \quad (1)$$

In Equation (1), ω is the laser modulation frequency, c the speed of sound, γ the Laplace coefficient of the gas, $k = \omega/c$ the wave number and \mathcal{H} the Fourier transform of the power density of the heat source. This pressure acoustic model is computationally efficient and accurate at the macro-scale but fails at the micro-scale (Glière et al 2014). In fact, various volume and surface dissipation processes, at work in the bulk of the propagation medium and close to the walls, cannot be neglected in miniaturized devices, where boundary layers occupy a

non-negligible part of the overall cell volume. Numerous approximate models have been adapted from the pressure acoustic model to take into account the dissipation effects. For instance, Kreuzer (1977) rely on eigenmode expansion of the pressure field and a correction by quality factors. The latter model is fast and faithful enough to constitute the coarse approximation used in our multifidelity approach. On the other hand, the high fidelity, but CPU time and memory consuming model relies on the full linearized Navier-Stokes formulation (FLNS), that accounts for viscous and thermal dissipation effects. In that approach, small harmonic variations are assumed to linearize the Navier-Stokes equation. The PDE equations system (2) is composed of the continuity equation, incorporating the ideal gas equation of state $p = \rho R_M T$, and the momentum and energy conservation laws:

$$\begin{aligned} i\omega \left(\frac{\tilde{p}}{p_0} - \frac{\tilde{T}}{T_0} \right) + \nabla \cdot \tilde{\mathbf{u}} &= 0 \\ i\omega \rho_0 C_p \tilde{T} &= -\nabla \cdot (-\kappa \nabla \tilde{T}) + i\omega \tilde{p} + Q_h \\ i\omega \rho_0 \tilde{\mathbf{u}} &= -\nabla \tilde{p} \\ &+ \nabla \cdot (\mu (\nabla \tilde{\mathbf{u}} + \nabla \tilde{\mathbf{u}}^T)) \\ &+ \nabla \cdot \left(\left(\lambda - \frac{2\mu}{3} \right) (\nabla \cdot \tilde{\mathbf{u}}) \mathbf{I} \right) \end{aligned} \quad (2)$$

where \tilde{p} , \tilde{T} and $\tilde{\mathbf{u}}$ are respectively the pressure, temperature and velocity fields in the gas, p_0 , T_0 and ρ_0 are the mean values of the pressure, temperature and density fields, λ and μ are the bulk viscosity and the shear viscosity. Q_h is the heat source.

2.2 Multi-fidelity metamodeling and optimization problem

In the framework of the PA sensor test case, the FLNS model constitutes the high fidelity model and the Kreuzer model the coarse version. Both models are solved using the commercial software package COMSOL multiphysics (COMSOL AB, Sweden), based on the finite element method. The computational time for the FLNS model is around one hour and ten minutes on a twenty-core cluster node cadenced at 3 GHz. For the Kreuzer model, the computational time on the same computer is reported at 3 minutes.

The PA cell has a chamber diameter of 1 mm and three design parameters are involved: chamber length, diameter and length of capillaries. Parameters ranges are available in Table 1. The cell resonance frequency and the maximum photoacoustic signal detected are the two outputs. The resonance frequency vary between 1000 Hz and 10000 Hz and the signal is around 1 Pa. In order to optimize the efficiency of the gas sensor,

the signal must be maximized. The resonance frequency must lie in a range limited for the low frequency by the $1/f$ noise and for high frequency, by the inverse of the molecules relaxation time.

Input	Range (mm)
chamber length	[10,20]
capillaries diameter	[0.3,1]
capillaries length	[5,20]

Table 1 Photoacoustic cell parameters and ranges

2.3 Single-fidelity metamodels comparison

A first study has been initiated to compare the prediction accuracy of Kriging and RBF metamodel on the photoacoustic gas sensor models. The purpose was to select the best metamodel for our test case and use it in an optimization sequence. The high fidelity and coarse model have been evaluated using a latin hypercube sampling maximin optimized of respectively 20 points and 100 points. The signal and the resonance frequency are surrogated and the prediction accuracy of metamodels are assessed using 10 test points. The prediction error is computed as the root-mean-squared error between predicted values and the output from the numerical models. Results in Table 2 are average over 5 different training sample.

Output	Model	High fidelity model	Coarse model
Resonance frequency	RBF	41.7±18.7	5.6±2.1
	Kriging	242.7±62.3	35.2±14.7
Signal	RBF	0.08±0.09	$9e^{-4} \pm 8e^{-4}$
	Kriging	0.06±0.02	0.005±0.001

Table 2 Prediction error comparison on both output over 5 initial designs

The prediction error for the RBF metamodel is lower than Kriging on 3 out of 4 cases. The signal approximated with the high fidelity dataset is the only case where the Kriging is superior in terms of accuracy. This result motivates the development of a RBF-based multifidelity metamodel, described in the next sections, to better address the photoacoustic cell optimization problem. **The difference in prediction accuracy between the two surrogate models is probably linked to the difference in parameter estimation methods : leave one out for RBF (see section 3) and maximum likelihood estimation for Kriging. This analysis is carried on in section 6.2.**

3 Radial Basis Function

Before explaining the multifidelity metamodel proposed in this work, the analytical basis of Radial Basis Function surrogate model is first recalled. We consider a real-valued function, formally defined by :

$$Y : \begin{cases} \mathcal{X} \subset \mathbb{R}^d \rightarrow \mathbb{R}^m \\ \mathbf{x} \mapsto \mathbf{z} = y(\mathbf{x}) \end{cases} \quad (3)$$

A surrogate model is built from a set of n input vectors $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ and a vector of corresponding scalar evaluations $\mathbf{z} = \{z_1, z_2, \dots, z_n\}$. The dimension of the input space is d . Multiple outputs are approximated one by one. Once the training of the surrogate model is achieved, **the prediction \hat{y} of the process output is obtained at new sample points**, with a highly reduced computational cost. The RBF is defined as a linear combination of basis functions φ that depend on the distance between the training points and the evaluation point:

$$\hat{Y}(\mathbf{x}) = \sum_{i=1}^n \beta_i \varphi(\|\mathbf{x} - \mathbf{x}_i\|) + Q(\mathbf{x}) \quad (4)$$

where $\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^d , the β_i coefficients are real numbers, Q is in the linear space Π_l of polynomials of degree at most l in \mathbb{R}^d and **the hat denotes the prediction value of the function**. It has been initially developed for scattered multivariate data interpolation by Dyn et al (1986). The polynomial is given by the following general form where \hat{l} is the polynomial degree, α is a vector of real coefficients and $p_k(\mathbf{x})$ are the monomial components:

$$Q(\mathbf{x}) = \sum_{k=1}^{\hat{l}} \alpha_k p_k(\mathbf{x}) \quad (5)$$

The basis function, the RBF coefficients β and the polynomial parameters (degree and coefficients) must be selected in order to build the RBF approximation. The interpolation condition, $\hat{Y}(\mathcal{X}) = \mathbf{z}$, leads to the linear system of equation, $\mathbf{z} = \Phi\beta$, where:

$$\Phi_{ij} := \varphi(\|\mathbf{x}_i - \mathbf{x}_j\|) \quad i, j = 1, \dots, n \quad (6)$$

The linear system is not sufficient to completely define the RBF. A solution proposed by Micchelli (1986) is to ensure that the basis function matrix Φ is conditionally positive definite, which reduces numerical errors when solving the system. The following inequality is obtained, where the polynomial order l_0 depends on the basis function:

$$(-1)^{l_0+1} \beta^T \Phi \beta > 0 \quad \forall \beta \in \mathcal{V}_{l_0} \setminus \{0\} \quad (7)$$

where $\mathcal{V}_{l_0} \subset \mathbb{R}^n$ is the linear space containing all $\beta \in \mathbb{R}^n$, that satisfy:

$$\sum_{i=1}^n \beta_i Q(\mathbf{x}_i) = 0 \quad \forall Q \in \Pi_{l_0} \quad (8)$$

The RBF φ used in this work and the corresponding polynomial form are given in Table 3, **where $\mathbf{x}[i]$ is the i -th component of \mathbf{x}** . The polynomial order l_0 is set at 1 for the Cubic case and -1 for the Gaussian one to fulfill the inequality in Equation (7). For a more complete list of RBF, the reader is referred to Gutmann (2001).

RBF	$\varphi(\ \mathbf{x} - \mathbf{x}'\)$	$Q(\mathbf{x})$
Cubic	$(\ \mathbf{x} - \mathbf{x}'\)^3$	$\alpha_1 + \alpha_2 \mathbf{x}$
Gaussian	$\exp\left(-\sum_{i=1}^d \gamma_i \ \mathbf{x}[i] - \mathbf{x}'[i]\ ^2\right), \gamma_i > 0$	$\{\emptyset\}$

Table 3 Different choices of radial basis functions

The unknown parameters β and α are now completely determined by the interpolation condition $\hat{Y}(\mathcal{X}) = \mathbf{z}$ and the positive definite condition in Equation (8) for the basis function matrix. Then the following system allows to estimate these parameters:

$$\begin{pmatrix} \Phi & \mathbf{F} \\ \mathbf{F}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \begin{pmatrix} \mathbf{z} \\ \mathbf{0} \end{pmatrix} \quad (9)$$

where $\mathbf{F} = (p_1(\mathcal{X}), \dots, p_{l_0+1}(\mathcal{X}))$. If the rank of matrix \mathbf{F} is equal to $d+1$, the matrix system is nonsingular and the resulting RBF interpolant is unique. The essential step of the unknown parameter estimation is the inversion of the nonsingular matrix. If the dataset is large, this is the most time demanding part of the surrogate model training. It is also noteworthy that the matrix is ill-conditioned if points in \mathcal{X} are close to each other.

Beyond determining the vector of basic parameters $(\beta^T \alpha^T)$, the parameter γ introduced in the case of Gaussian RBF must also be estimated. The correct estimation of this parameter enables to minimize the generalization error of the model. It can be interpreted as a scaling factor which expresses the spacial influence in each direction of the basis function central point. Forrester et al (2008) suggests to use the cross-validation error estimate in order to determine the value of γ . First, γ is selected from candidate values and then basic parameters are determined for each subset of the cross-validation population. Once a value with the minimal cross-validation error is found, basic parameters are built on the whole training dataset with the optimized γ . Rippa (1999) derived a leave-one-out (LOO) formula (Equation (10)) that estimates the model error

by determining basic parameter only once by γ value. This method is used in the present work.

$$LOO = \sum_{i=1}^n \left(\frac{\beta_i}{(\Phi^{-1})_{ii}} \right)^2 \quad (10)$$

The LOO criterion optimization was solved using the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) (Hansen and Kern 2004). CMA-ES is a second-order evolutionary search strategy based on the propagation of a covariance matrix, which has produced excellent results on multimodal and high dimensional test problems.

Now the prediction capability is studied. A standard uncertainty measure is derived by Jakobsson et al (2009), inspired by the interpretation of the native space norm as a “bumpiness” measure by Gutmann (2001). This quantity s_Y^2 , also called power function, describes the interpolation error at a point \mathbf{x} of the input space:

$$s_Y^2(\mathbf{x}) = (-1)^{l_0+1} \times \left[\varphi(0) - (\mathbf{u}_x^T \mathbf{F}_x^T) \left(\begin{pmatrix} \Phi & \mathbf{F} \\ \mathbf{F}^T & \mathbf{0} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{u}_x \\ \mathbf{F}_x \end{pmatrix} \right) \right] \quad (11)$$

where $\mathbf{u}_x = (\varphi(\|\mathbf{x} - \mathbf{x}_1\|), \dots, \varphi(\|\mathbf{x} - \mathbf{x}_n\|))^T$ and $\mathbf{F}_x = (p_1(\mathbf{x}), \dots, p_{l_0+1}(\mathbf{x}))$. An example of Cubic and Gaussian RBF interpolant is available on Figure 2, together with the bumpiness measure of Gutmann. For both basis functions, the prediction error measure indicates to add a new point at $\mathbf{x} = 1$ in order to reduce the uncertainty in this area. This prediction error has also large values in area where training points are sparse, for instance close to $\mathbf{x} = 0.4$.

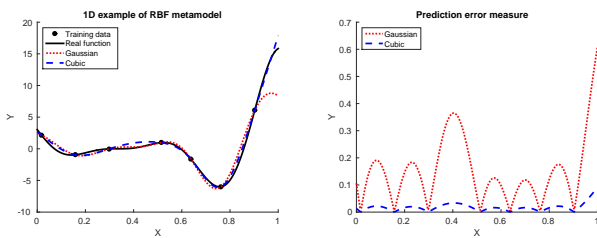


Fig. 2 Example of RBF interpolation and associated bumpiness used as prediction error measure

Even though the Kriging model framework is not detailed here, it is important to recall that RBF is similar to the Kriging model, as stated by Costa et al (1999) and Mackman et al (2013). The basis function and the polynomial term can be the same for both metamodels,

which lead to identical interpolants if the parameters are estimated with the same method. The difference between these formulations is more conceptual, since they were derived from different branches of mathematics. For example, Kriging assumes that the process Y to surrogate is the realization of a Gaussian process (Santner et al 2003), which is not the case for RBF.

4 Co-RBF

Designing a component such as a mechanical structure or optical planar integrated circuit may involve a computationally-demanding and detailed numerical simulation. In many cases, it is feasible to obtain a simpler version of the model using approximations that is qualitatively similar to the full model. For example, Vitali et al (2002) used a finite element model of a stiffened panel with a crack for the high fidelity model to capture accurately the stress next to the crack tip, and a low fidelity model without the crack to compute nominal stresses and strains. Here, it is assumed that different levels of modeling, in terms of accuracy, are available for our physical problem. It is possible to generalize the approach to multiple sets of data with different fidelity levels but, for readability sake, the present method description is limited to two datasets. The most accurate set, which is also the more demanding in terms of computational time, is represented by training points \mathcal{X}_e and their corresponding outputs \mathbf{z}_e . The coarse version of the simulation model is represented by \mathcal{X}_c and \mathbf{z}_c . It is assumed that the high fidelity sampling points are also in the training dataset of the low fidelity model. With this condition, the multifidelity approximation is built using the auto-regressive model of Kennedy and O’Hagan (2000) :

$$\hat{y}_e(\mathbf{x}) = \rho \hat{y}_c(\mathbf{x}) + \hat{y}_d(\mathbf{x}) \quad (12)$$

where ρ is a scaling factor and y_d the function representing the difference between the expensive process and the scaled coarse process. This formulation allows different levels of fidelity to have different correlation structures. This implies that the RBF parameters for the coarse process are not be the same as the parameters for the expensive one. In order to estimate the RBF of the coarse process, the coarse dataset is considered independently of the expensive one and this leads to the parameters estimation described by Equation (4). Once $(\beta_c^T \alpha_c^T)$ are obtained, the RBF parameters related to the difference function have to be estimated. With the auto-regressive model, the assumption is made that, given the point \mathbf{x}_i evaluated on the coarse model, no more information can be learnt about \mathbf{z}_{e_i}

from the coarse model. This indicates that only $y_c(\mathbf{x}_e)$ is considered in this estimation. If the coarse process has not already been evaluated at these points, the values predicted by the previously built RBF are used, which relax the condition that $\mathcal{X}_e \subset \mathcal{X}_c$. Then the following linear system has to be solved, where $(\beta_d^T \alpha_d^T)$ corresponds to the RBF parameters of the difference model:

$$\begin{pmatrix} \Phi_d & \mathbf{F}_e \\ \mathbf{F}_e^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \beta_d \\ \alpha_d \end{pmatrix} = \begin{pmatrix} \mathbf{z}_e - \rho \hat{y}_c(\mathbf{x}_e) \\ \mathbf{0} \end{pmatrix} \quad (13)$$

This linear system depends on the scaling factor ρ . As for the Gaussian RBF parameter, the scaling factor may be determined using the leave-one-out error estimation on the expensive dataset. If both γ and ρ have to be estimated, the minimization of the LOO error has to be done regarding both parameters. Once optimal parameters have been determined, the surrogate model of the expensive process is computed using Equation (12).

An example of co-RBF metamodel is available on Figure 3 left, in an analytical test case defined by Forrester et al (2007). Four expensive points and eleven coarse points are evaluated to build the Gaussian co-RBF. The resulting interpolant is the dashed line and it is close the real model (solid line). The evolution of the LOO depending on γ_d and ρ is also plotted on the right. The global minimum of the map is located at $\gamma_d = 1$ and $\rho = 1.74$.

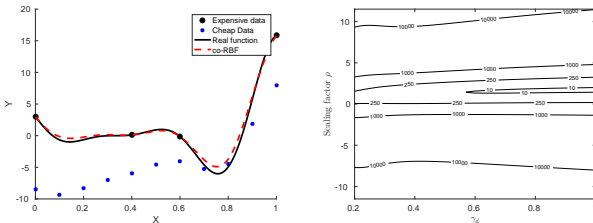


Fig. 3 LOO cross-validation error on a one-dimensional example of co-RBF metamodeling.

5 Co-RBF Based Optimization

Gutmann (2001) developed an adaptive sampling method for optimization in a single fidelity RBF approach, which is similar to the Efficient Global Optimization (EGO) method of Jones et al (1998). It is based on the standard uncertainty measure of Equation (11) and an estimate of the real function minimum value. This algorithm has then been modified by Regis

and Shoemaker (2006) to improve its local search property. Here the so-called original Gutmann-RBF algorithm is derived for the multifidelity RBF framework. The estimated global minimum of the high fidelity function is denoted as z^* . The co-RBF interpolation is rebuilt with the modified accurate training set containing $(\mathcal{X}_e, \mathbf{z}_e)$ and (x_e^*, z^*) where $x_e^* \notin \mathcal{X}_e$. The point x_e^* to be evaluated and added to training dataset, is the one that leads to a re-built co-RBF which is most "reasonable". An example of what reasonable means is available on Figure 4 with a single fidelity approach. The target value chosen is represented by the dashed line and two different locations are tested. The upper right plot is the negative log value of the Gutmann criterion. The lower right plot is the best infill points to evaluate than the one on the lower left plot according to this criterion. This point is also in the area of the true global optimum of the real function.

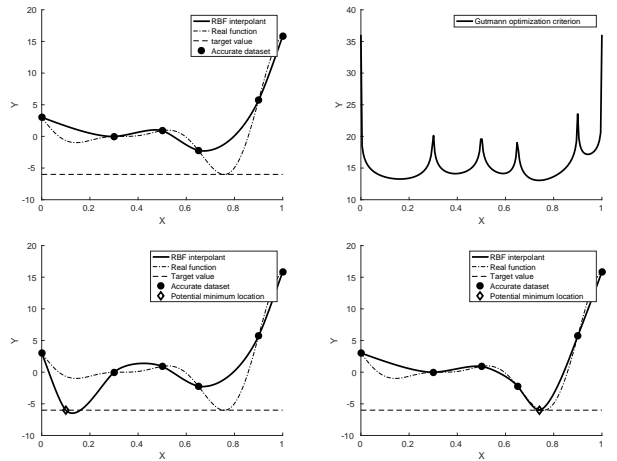


Fig. 4 Example of next evaluation point selection using Gutman-RBF algorithm. Upper left plot is the actual state of the RBF. Upper right is negative log value of Gutmann criterion to minimize. Lower plot are two examples of RBF re-built with a target value point.

Gutmann derived an infill criterion that indicates the next evaluation point to be added to the RBF metamodel training dataset. After multiple iterations, the minimum of the RBF converges toward the global minimum of the function. The following assumptions are needed to transpose this method into the co-RBF framework. First, the evaluation time of the coarse model is considered negligible compare to the accurate model, then it can be evaluated a large number of times. Thus, it is possible to build an accurate metamodel of this function and there is no need to add points to the coarse training dataset during optimization. The expensive function is the only one that is called during the

refinement process. Second, since the difference model is the one built using accurate model evaluation, the infill criterion of Gutmann is computed using the prediction error of RBF model φ_d . The formula derived in Equation (14) is the one used in the multifidelity framework:

$$h_n(\mathbf{x}) = \frac{(-1)^{l_0+1}}{[\hat{y}_e(\mathbf{x}) - z^*]^2} \times \left[\varphi_d(0) - (u_n^T \mathbf{F}_x^T) \left(\begin{bmatrix} \Phi_d & \mathbf{F}_d \\ \mathbf{F}_d^T & \mathbf{0} \end{bmatrix}^{-1} \begin{pmatrix} u_n \\ \mathbf{F}_x \end{pmatrix} \right) \right] \quad (14)$$

where $u_n(\mathbf{x}) := (\varphi(\|\mathbf{x} - \mathbf{x}_1\|), \dots, \varphi(\|\mathbf{x} - \mathbf{x}_n\|))^T$ and $\mathbf{F}_x = [\mathbf{x}^T \ 1]$. The selection of the goal value z^* for the global optimum has to be explained. In Gutmann approach, it is cyclically defined in the interval $[-\infty, \min(\hat{y}_e)]$, where low values corresponds to global search while values close to the actual minimum of the metamodel leads to local search. The convergence of the whole optimization algorithm to the global minimum is proven (Gutmann 2001) in the linear, cubic and thin plate spline cases if the goal value is selected using the following procedure. Let N be the cycle length and p be a permutation of $\{1, \dots, n\}$ such that $y(x_{p(1)}) \leq \dots \leq y(x_{p(n)})$. The number of points in the initial training set is n_0 and n is the current number of points in the training dataset during optimization. The goal value is set as follow:

$$z^* = \min \hat{y}_e - W_n \cdot (y(x_{p(k_n)}) - \min \hat{y}_e) \quad (15)$$

where

$$W_n = \left[\frac{\text{mod}(N - (n - n_0), N + 1)}{N} \right]^2$$

$$k_n = \begin{cases} n, & \text{if } \text{mod}(n - n_0, N + 1) = 0 \\ k_{n-1} - \frac{(n - n_0)}{N}, & \text{otherwise} \end{cases} \quad (16)$$

The values of W_n are decreasing during a cycle so that it allows moving from a global to a local infill criterion. At each iteration, the quantity $-\log(h_n)$ is minimized in order to avoid numerical difficulties. Once again, the CMA-ES algorithm is used to get the optimum of the Gutmann criterion since it is a multi-modal objective function. The cycle length is set to 5 in this work.

6 Analytical Benchmark

A benchmark of two analytical metamodeling test problems, with different characteristics, has been defined to compare multifidelity surrogate modeling methods. The

problems have been used by Xiong et al (2013) for illustration on multifidelity sequential design. The first one is a simple two-dimensional problem where the basic co-RBF prediction capability is tested. The second problem allows checking if higher dimensional problems are better approximated by co-RBF or by co-kriging. The co-kriging model is built using the ooDACE Toolbox (Couckuyt et al 2014) and the co-RBF code is implemented using the commercial software package MATLAB (The MathWorks Inc., U.S.A.). For each problem, the co-RBF prediction error is compared to the co-kriging one. The accuracy is analyzed on different sampling plan sizes and results are averaged on one hundred initial random space-filling designs. The evolution of the prediction error over the combination of different expensive and coarse evaluations is plotted. The root-mean-squared error (RMSE) on one hundred test points of the expensive code is computed for both metamodels as the measure of accuracy.

6.1 Test function 1

The first test problem is based on a function from Currin et al (1991). The high fidelity function and the coarse version are detailed in Equation (17).

$$y_e = \left[1 - \exp\left(-\frac{1}{2x_{(2)}}\right) \right] \frac{2300x_{(1)}^3 + 1900x_{(1)}^2 + 2092x_{(1)} + 60}{100x_{(1)}^3 + 500x_{(1)}^2 + 4x_{(1)} + 20}$$

$$y_c = (1/4) \times [y_e(x_{(1)} + 1/20, x_{(2)} + 1/20) + y_e(x_{(1)} + 1/20, \max(0, x_{(2)} - 1/20)) + y_e(x_{(1)} - 1/20, x_{(2)} + 1/20) + y_e(x_{(1)} - 1/20, \max(0, x_{(2)} - 1/20))] \quad (17)$$

The result obtained for this problem are globally the same for both methods. **The average of the mean prediction error on Figure 5 is 0.26 ± 0.12 for co-kriging and 0.22 ± 0.10 for co-RBF.** As expected, if the number of expensive or coarse model evaluations in the training sample increases, the accuracy of both models is improved. Some data among the hundred initial designs are not available for co-kriging because the ooDACE toolbox includes a protection mechanism against too ill-conditioned correlation matrix. The difference on this example is on the standard deviation of the RMSE (Figure 6), which is slightly more important for co-kriging than co-RBF for low expensive model evaluation numbers.

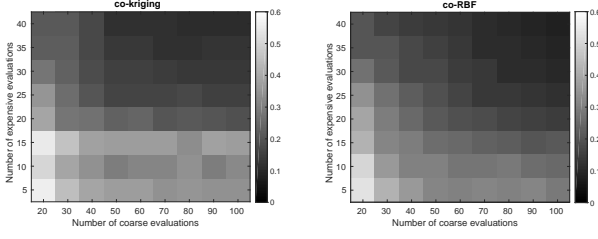


Fig. 5 Mean value of the prediction error on Currin et al (1991) problem over 100 random initial designs

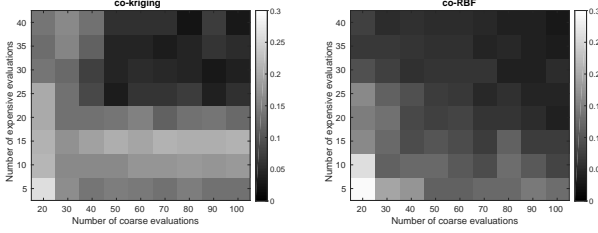


Fig. 6 Standard deviation of the prediction error on Currin et al (1991) problem over 100 random initial designs

6.2 Test function 2

The second test problem is based on the borehole model (Morris et al 1993) and is derived by Xiong et al (2013) as a multi-fidelity model. It describes the flow of water through a borehole drilled from the ground surface through two aquifers. The eight inputs and their ranges are detailed in Table 4. Cubic co-RBF and Gaussian co-RBF are also compared on this larger dimensional example described in Equation (18).

$$y_e = \frac{2\pi T_u (H_u - H_l)}{\log(r/r_w) \left[1 + \frac{2LT_u}{\log(r/r_w)r_w^2 K_w} + T_u/T_l \right]} \quad (18)$$

$$y_c = \frac{5\pi T_u (H_u - H_l)}{\log(r/r_w) \left[1.5 + \frac{2LT_u}{\log(r/r_w)r_w^2 K_w} + T_u/T_l \right]}$$

Input	Range
r_w : radius of borehole	[0.05, 0.15]
r : radius of influence	[100, 50000]
T_u : transmissivity of upper aquifer	[63070, 115600]
H_u : potentiometric head of upper aquifer	[990, 1110]
T_l : transmissivity of lower aquifer	[63.1, 116]
H_l : potentiometric head of lower aquifer	[700, 820]
L : length of borehole	[1120, 1680]
K_w : hydraulic conductivity of borehole	[9855, 12045]

Table 4 Inputs and ranges of the borehole model

The mean value of the prediction error with each multi-fidelity metamodel is plotted on Figure 7. The scaling

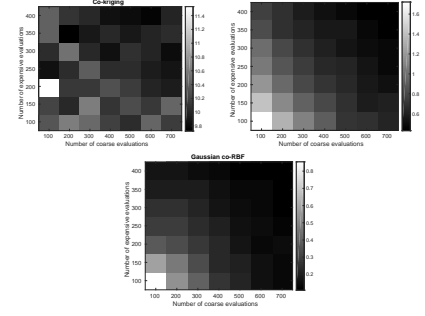


Fig. 7 Mean value of the prediction error on the borehole problem over 10 random initial designs

is different for each figure in order to better distinguish the differences in trend between accuracy and number of evaluations points. The average of the mean prediction error is 10.1 ± 0.3 for co-kriging, 0.7 ± 0.3 for Cubic co-RBF and 0.2 ± 0.1 for the Gaussian one. Both Co-RBF models are superior to co-kriging on this example, and their accuracy increases when the number of expensive and coarse evaluation increases. Co-kriging does not approximate correctly the borehole model because its prediction error is only slightly less than 10%. For co-RBF, as expected, the Gaussian version is more accurate than the cubic one. The parameter γ of the Gaussian co-RBF version has a different value for each dimension of the borehole model and thus provides a more accurate approximation of the problem. The standard deviation of the error on Figure 8 reveals the same behavior as for the mean value.

The random pattern noticeable on the co-kriging re-

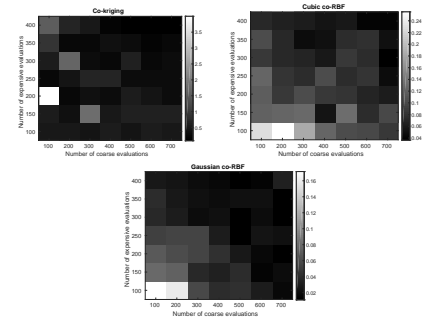


Fig. 8 Standard deviation of the prediction error on the borehole problem over 10 random initial designs

sults could be explained by the metamodel parameters estimation. The maximum likelihood estimation is often difficult to solve when the number of parameters increases and it is the method used here for co-kriging. The leave-one-out procedure used for co-RBF builds an interpolant that is less sensitive to a loss of information

in the training dataset. It also seems to bring more stable results among the different sampling plan.

7 Metamodeling and Optimization of the photoacoustic cell

Multi-fidelity metamodels are built, with Gaussian basis functions, using five to twenty expensive function calls and ten to one hundred coarse function calls. Each combination of expensive and coarse evaluations are built over five different initial samples. The prediction accuracy is computed using the RMSE on the same ten extra points of the expensive code for all results. The RMSE is then averaged over the five different results by combination of expensive and coarse model.

The mean prediction error of the approximation of the cell resonance frequency is plotted on Figure 9. Here co-RBF brings a better approximation than co-kriging for each combination and the error on the prediction value is less than 1% at the best combination. As a quantitative measure, the average of the mean prediction error of the co-kriging is 91 ± 56 and the one for co-RBF is 28 ± 14 . The accuracy of co-RBF is improved when the number of coarse and expensive evaluations increases but this tendency is less clear for co-kriging. The standard deviation is plotted on Figure 10. It appears also that co-RBF provides less scattered result than co-kriging.

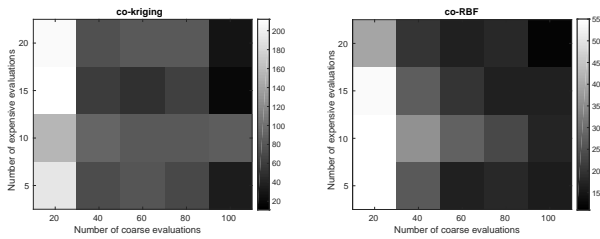


Fig. 9 Mean value of the prediction error on cell resonance frequency over 5 random initial designs

The second output, i.e. the maximum signal detected by the PA cell, is now approximated. The mean prediction error is presented on Figure 11. The co-kriging gives better results than co-RBF on the overall combination map but there is not a clear correlation between more accuracy and larger number of evaluations points. The average of the mean prediction error of the co-kriging is 0.07 ± 0.06 and the one for co-RBF is 0.07 ± 0.01 . Moreover, the standard deviation plot of Figure 12 shows that RMSE values for each initial

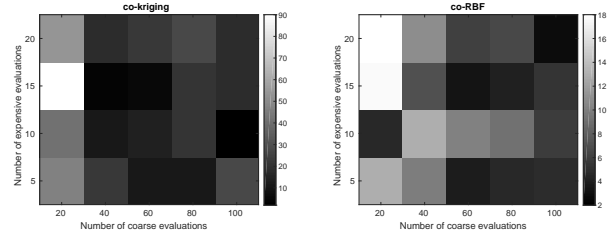


Fig. 10 Standard deviation of the prediction error on cell resonance frequency over 5 random initial designs

design are scattered for both methods. The FLNS model may model easily the acoustic signal of the PA cell. Thus, it may not be useful to use the Kreuzer model in a multifidelity metamodeling of the signal. The number of expensive points has to be increased in order to improve significantly the prediction accuracy.

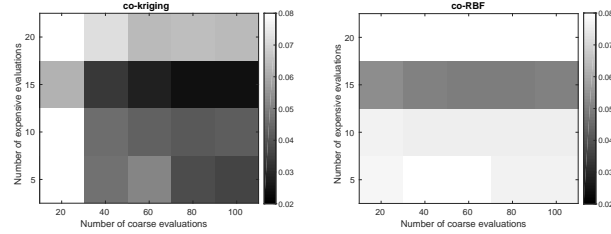


Fig. 11 Mean value of the prediction error on cell signal over 5 initial designs

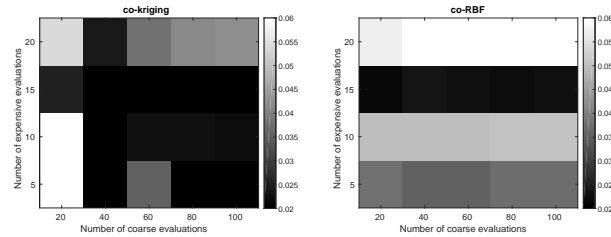


Fig. 12 Standard deviation of the prediction error on cell signal over 5 initial designs

In order to assess the benefit of using a multi-fidelity approach for the approximation of both outputs, single-fidelity and multi-fidelity RBF metamodels are compared. Table 5 summarizes the prediction error when RBF model is built on the same expensive datasets that are used for the multi-fidelity problem. The table includes a reminder of co-RBF and co-kriging accuracy obtained with 100 coarse evaluations. For each number of expensive calls, the results are once again averaged over the same five different initial

Output	Model	5 points	10 points	15 points	20 points
Resonance frequency	RBF	223±41	147±75	63±38	42±19
	co-RBF	16±5	18±6	17±5	12±3
	co-kriging	33±27	86±2	20±18	28±18
Signal	RBF	0.12±0.07	0.09±0.06	0.07±0.04	0.08±0.09
	co-RBF	0.08±0.04	0.08±0.05	0.05±0.02	0.08±0.06
	co-kriging	0.04±0.01	0.04±0.02	0.02±0.06	0.06±0.04

Table 5 Prediction error comparison on both output over 5 initial designs of expensive evaluations. Multi-fidelity metamodels results are those obtained with 100 coarse evaluations.

samples.

The accuracy on the resonance frequency prediction is clearly improved with multi-fidelity modeling compared to the single fidelity one. The RBF prediction leads to results similar to those of the multi-fidelity method, as the number of expensive evaluations increases. Results on the detected signal show that multifidelity and classic RBF prediction errors are not very different. The Kreuzer model for the signal adds information to multifidelity model but it does not improve accuracy as much as for the resonance frequency output.

Here, the improvement brought by the multifidelity approach on the prediction accuracy of the resonance frequency can be used to reduce the computational time of the FLNS model. Since the result brought by this time consuming model depends on a frequency sweep, it is possible to reduce the search range if an estimate of the cell resonance frequency is available. The actual sweep is 50 values between 1000 Hz and 10000 Hz, then the resonance is estimated by spline interpolation of the dataset. By using the prediction of the resonance frequency of the co-RBF, the sweep range is reduced to 5 values and thus, the computational time of the FLNS to 40 minutes. This strategy is now implemented when the FLNS model is called and a co-RBF on the resonance frequency has already been built (for example during a sequential optimization of the photoacoustic cell using co-RBF).

7.1 Optimization of the gas sensor geometry

Since a multifidelity metamodel of the signal detected by the sensor is now available, an adaptative sampling strategy to get the optimum value is set up. The method described in Section 5 is used as an infill criterion for the co-RBF. It is compared to the traditional Efficient Global Optimization (EGO) algorithm adjusted to the multifidelity framework (Forrester et al 2007). Five different initial DOE are tested with 100 points from the coarse model and 5 from the accurate one. The number of additional calls to the accurate simulation model is arbitrarily set to 16. The

results are presented in Table 6, where the value of the minimum obtained and its location are compared for each method.

Both methods lead to similar results. Two optimal locations are found, the first one when the capillaries and the chambers length are equal and the second when the capillaries and chambers diameter are equal. No significant difference in sensor performance is noted in the two solutions obtained.

The photoacoustic cell could be approximated by a spring-mass system, where capillaries are the mass and chambers the springs. The two optimal solutions give a close equivalent mass, which may explain that the sensor behavior is nearly the same for both geometries. Optimums lie close to the parameters space boundaries, but it is not possible to relax them, as the bounds have been chosen to fulfill some integration requirements of the gas sensor in a higher level system. In terms of algorithm performances, Gutmann method adapted to multifidelity produced equivalent results to the well known EGO and somewhat more accurate minimum location. This first comparison would interestingly be extended to a higher dimensional problem, where the co-RBF seems to be more accurate.

8 Conclusion

A methodology for multifidelity surrogate model based on RBF has been proposed. The metamodel is constructed using an auto-regressive model that was first introduced for co-kriging. A comparison on an analytical benchmark shows that co-kriging and co-RBF bring similar results on low dimensional problems. In higher dimension, co-RBF outperforms co-kriging on our test functions. The Gaussian version of co-RBF appears to be more efficient on multidimensional problem than the cubic version. This first analysis suggests to consider co-RBF as an alternative to co-kriging in high dimensional problems. The new method has also been validated on a real engineering problem: the design of a photoacoustic based gas sensor. The benefit of a multifidelity approach to approximate the resonance frequency of the cell has

Initial DOE	Minimum value		Minimum location	
	Co-kriging	Co-RBF	Co-kriging	Co-RBF
1	-1.88	-1.91	[0.02, 4.18e ⁻⁴ , 0.0179]	[0.02, 5.00e ⁻⁴ , 0.0177]
2	-1.87	-1.90	[0.02, 5.00e ⁻⁴ , 0.0159]	[0.02, 4.97e ⁻⁴ , 0.0173]
3	-1.91	-1.90	[0.02, 4.86e ⁻⁴ , 0.0188]	[0.02, 4.73e ⁻⁴ , 0.0200]
4	-1.90	-1.88	[0.02, 4.54e ⁻⁴ , 0.0200]	[0.02, 5.00e ⁻⁴ , 0.0200]
5	-1.87	-1.91	[0.02, 4.58e ⁻⁴ , 0.0200]	[0.02, 4.46e ⁻⁴ , 0.0200]

Table 6 Comparison between co-kriging and co-RBF based optimization on the photoacoustic cell design

been demonstrated. Then an extension of the Gutmann algorithm to multifidelity framework has been applied to optimize the photoacoustic cell. A benchmark between co-kriging and co-RBF based optimization on the gas sensor design led to two optimal solutions. The efficiency of the co-RBF based method is close to that of EGO procedure, but performed somewhat better on optimum location. In the near future, these promising results will be completed by studying the optimization of photonic planar integrated circuits involving a larger number of parameters.

References

- Bauer R, Stewart G, Johnstone W, Boyd E, Lengden M (2014) "3D-printed miniature gas cell for photoacoustic spectroscopy of trace gases". *Optics Letters* 39(16):4796–4799, DOI 10.1364/OL.39.004796
- Costa JP, Pronzato L, Thierry E (1999) A comparison between Kriging and radial basis function networks for nonlinear prediction. In: *Nonlinear Signal and Image Processing*, pp 726–730
- Couckuyt I, Dhaene T, Demeester P (2014) "ooDACE toolbox: a flexible object-oriented kriging implementation". *The Journal of Machine Learning Research* 15(1):3183–3186
- Curran C, Mitchell T, Morris M, Ylvisaker D (1991) "Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments". *Journal of the American Statistical Association* 86(416):953, DOI 10.2307/2290511
- Dong H, Song B, Wang P, Huang S (2015) Multi-fidelity information fusion based on prediction of kriging. *Structural and Multidisciplinary Optimization* 51(6):1267–1280, DOI 10.1007/s00158-014-1213-9
- Dyn N, Levin D, Rippa S (1986) "Numerical procedures for surface fitting of scattered data by radial functions". *SIAM Journal on Scientific and Statistical Computing* 7(2):639–659, DOI 10.1137/0907043
- Forrester AI, Söbester A, Keane AJ (2007) "Multi-fidelity optimization via surrogate modelling". *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 463(2088):3251–3269, DOI 10.1098/rspa.2007.1900
- Forrester AI, Söbester A, Keane AJ (2008) "Engineering design via surrogate modelling a practical guide". J. Wiley, Chichester, West Sussex, England; Hoboken, NJ
- Glière A, Rouxel J, Brun M, Parvitte B, Zéninari V, Nicoletti S (2014) "Challenges in the design and fabrication of a lab-on-a-chip photoacoustic gas sensor". *Sensors* 14(1):957–974, DOI 10.3390/s140100957
- Gutmann HM (2001) "A radial basis function method for global optimization". *Journal of Global Optimization*
- Hansen N, Kern S (2004) "Evaluating the CMA evolution strategy on multimodal test functions". In: Yao X, et al (eds) *Parallel Problem Solving from Nature PPSN VIII*, Springer, LNCS, vol 3242, pp 282–291
- Holthoff E, Heaps D, Pellegrino P (2010) "Development of a mems-scale photoacoustic chemical sensor using a quantum cascade laser". *Sensors Journal, IEEE* 10(3):572–577, DOI 10.1109/JSEN.2009.2038665
- Jakobsson S, Andersson B, Edelvik F (2009) Rational radial basis function interpolation with applications to antenna design. *Journal of Computational and Applied Mathematics* 233(4):889–904, DOI 10.1016/j.cam.2009.08.058
- Jin R, Chen W, Simpson T (2001) Comparative studies of metamodeling techniques under multiple modelling criteria. *Structural and Multidisciplinary Optimization* 23(1):1–13, DOI 10.1007/s00158-001-0160-4
- Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. *Journal of Global optimization* 13(4):455–492
- Kennedy MC, O'Hagan A (2000) "Predicting the output from a complex computer code when fast approximation are available". *Biometrika* 87(1):1–13
- Kreuzer LB (1977) "The physics of signal generation and detection". In: Pao YH (ed) *Optoacoustic Spectroscopy and Detection*, Academic Press, pp 1–25
- Krige DG (1951) "A statistical approach to some basic mine valuation problems on the witwatersrand". *Journal of the Chemical, Metallurgical and Mining Society of South Africa* 52(6):119–139
- Le Gratiet L, Cannamela C (2015) "Cokriging-Based sequential design strategies using fast cross-validation techniques for multi-fidelity computer codes. *Technometrics* 57(3):418–427, DOI 10.1080/00401706.2014.928233
- Mackman TJ, Allen CB, Ghoreyshi M, Badcock KJ (2013) Comparison of Adaptive Sampling Methods for Generation of Surrogate Aerodynamic Models. *AIAA Journal* 51(4):797–808, DOI 10.2514/1.J051607
- Micchelli CA (1986) Interpolation of scattered data: Distance matrices and conditionally positive definite functions. *Constructive Approximation* 2(1):11–22, DOI 10.1007/BF01893414
- Miklós A, Hess P, Bozóki Z (2001) "Application of acoustic resonators in photoacoustic trace gas analysis and metrology". *Review of Scientific Instruments* 72(4):1937–1955, DOI 10.1063/1.1353198
- Morris MD, Mitchell TJ, Ylvisaker D (1993) Bayesian design and analysis of computer experiments: Use of derivatives in surface prediction. *Technometrics* 35(3):243–255
- Powell MJD (1987) Radial basis functions for multivariable interpolation: a review. In: Mason JC, Cox MG (eds) *Algorithms for Approximation*, Clarendon Press, New York, NY, USA, pp 143–167

- Powell MJD (2001) "Radial basis function methods for interpolation to functions of many variables". In: HERCMA, pp 2–24
- Regis RG, Shoemaker CA (2006) Improved Strategies for Radial basis Function Methods for Global Optimization. *Journal of Global Optimization* 37(1):113–135, DOI 10.1007/s10898-006-9040-1
- Rippa S (1999) "An algorithm for selecting a good value for the parameter c in radial basis function interpolation". *Advances in Computational Mathematics* 11(2-3):193–210
- Rouxel J, Coutard JG, Gidon S, Lartigue O, Nicoletti S, Parvitte B, Vallon R, Zninari V, Glière A (2016) Miniaturized differential Helmholtz resonators for photoacoustic trace gas detection. *Sensors and Actuators B: Chemical* DOI 10.1016/j.snb.2016.06.074
- Santner TJ, Williams BJ, Notz WI (2003) "The Design and Analysis of Computer Experiments". Springer-Verlag, Berlin-Heidelberg
- Schonlau M, Welch WJ (1996) "Global optimization with nonparametric function fitting". *Proceedings of the ASA, Section on Physical and Engineering Sciences* pp 183–186
- Sun G, Li G, Stone M, Li Q (2010) A two-stage multi-fidelity optimization procedure for honeycomb-type cellular materials. *Computational Materials Science* 49(3):500–511, DOI 10.1016/j.commatsci.2010.05.041
- Vitali R, Haftka R, Sankar B (2002) Multi-fidelity design of stiffened composite panel with a crack. *Structural and Multidisciplinary Optimization* 23(5):347–356, DOI 10.1007/s00158-002-0195-1
- Wang GG, Shan S (2007) "Review of metamodeling techniques in support of engineering design optimization". *Journal of Mechanical Design* 129(4):370–380
- Xiong S, Qian PZG, Wu CFJ (2013) "Sequential design and analysis of high-accuracy and low-accuracy computer codes". *Technometrics* 55(1):37–46, DOI 10.1080/00401706.2012.723572
- Zéninari V, Kapitanov VA, Courtois D, Ponomarev YN (1999) "Design and characteristics of a differential helmholtz resonant photoacoustic cell for infrared gas detection". *Infrared physics & technology* 40(1):1–23